Supporting information

Coupling Isoelectric Focusing Gel Electrophoresis to Mass Spectrometry by Electrostatic Spray Ionisation

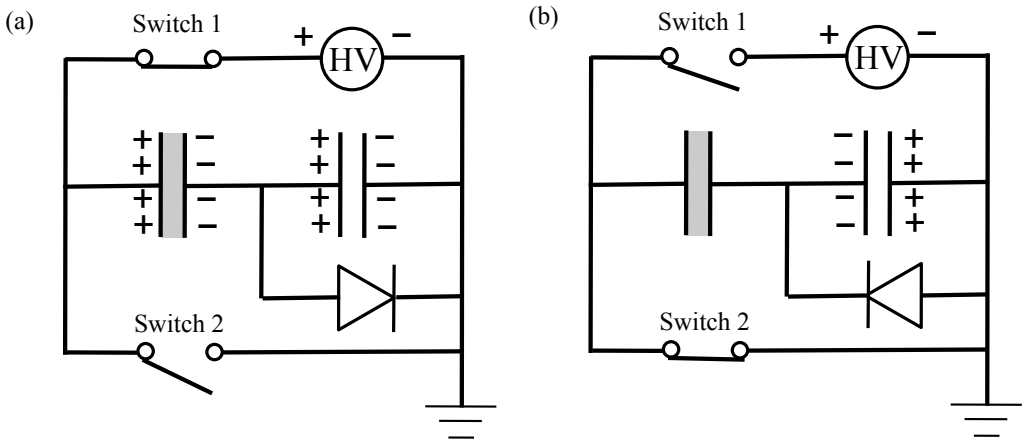
Liang Qiao‡, Elena Tobolkina‡, Baohong Liu and Hubert H. Girault*

‡: LQ and ET contributed equally to this work

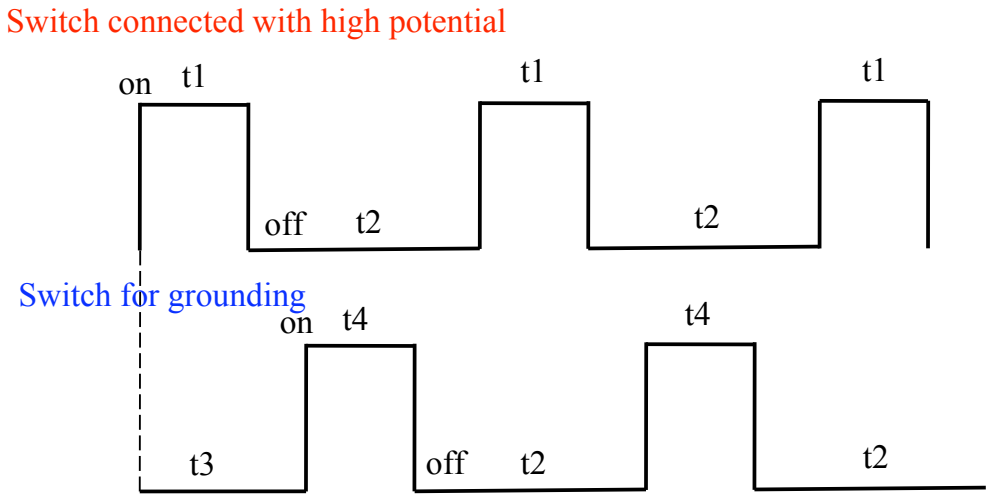
*: Correspondence should be addressed to HHG (hubert.girault@epfl.ch)

Table of content

SI-1: Electronic circuit and Labview programme to synchronize two switches



Scheme SI-1.1 Electronic circuit of electrostatic spray ionization during capacitors (a) charging and (b) discharging. The diode symbolises the direction of the spray current: spray of cations in (a) and spray of anions in (b).



Scheme SI-1.2 Schematic illustration of the performance of the two switches controlled by a Labview programme.

SI-2: ESTASI-MS detection of proteins in gel with the help of a plastic cover patterned with holes

The gel fixed on a piece of plastic plate patterned with holes was used for electrophoresis. Cytochrome c solution (1 µl, 0.2 mg/ml) was deposited on the gel strip. The electrophoresis was performed for 10 min (300 V, 1 mA) with an EPS 3501 XL power supply (Amersham Pharmacia Biotech, Sweden). Afterwards, the gel was washed briefly with water for ESTASI-MS analysis of proteins. Shown as figure SI-2, cytochrome c was detected; indicating that ESTASI-MS also worked when a plastic cover patterned with holes was fixed on the gel.

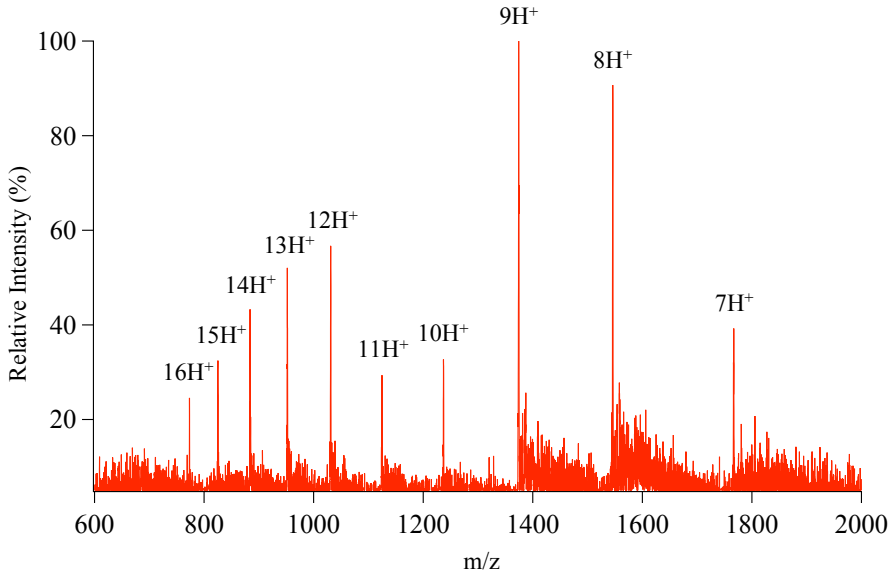


Figure SI-2. Mass spectrum of 200 ng cytochrome c in a polyacrylamide gel pH 4. The gel was under a plastic cover drilled with holes for filling acidic buffer. The ions were generated by ESTASI when a pulsed positive high potential (6.5 kV) was applied to the electrode. The label shows the charge states.

SI-3: Spatial resolution of gel ESTASI-MS

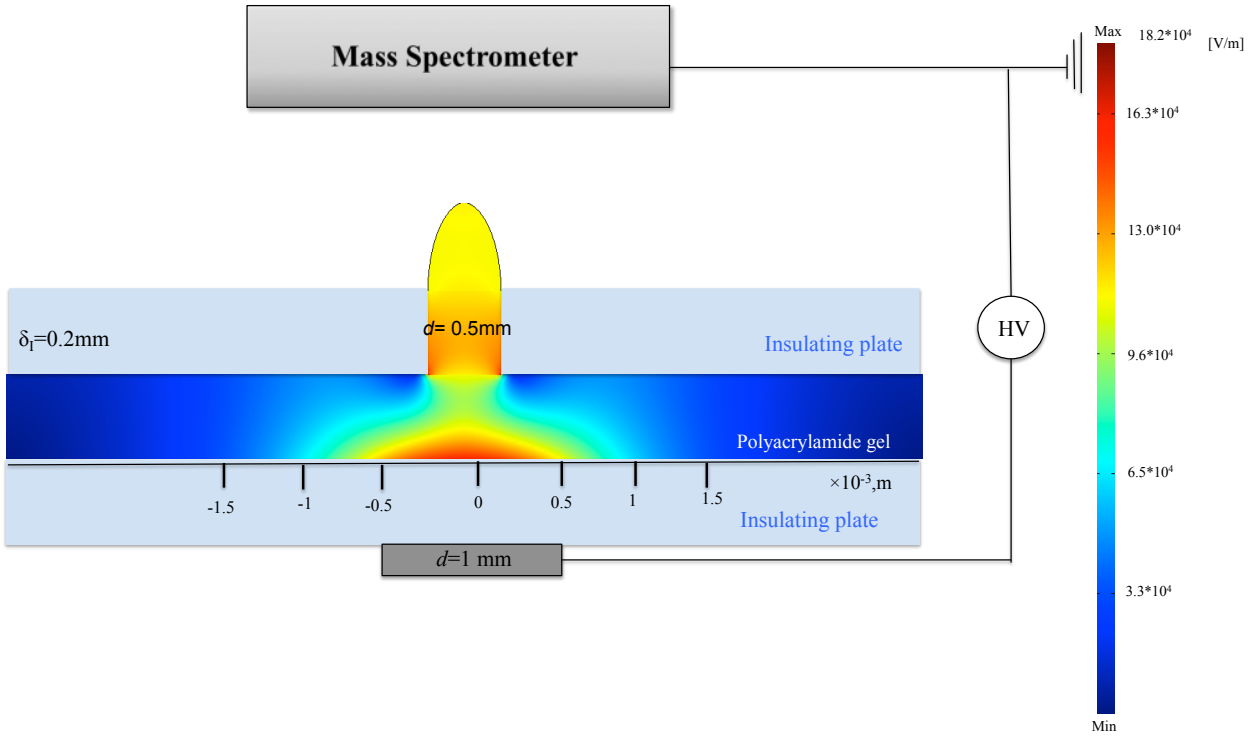


Figure SI-3. Finite element simulation of electric field in a polyacrylamide gel during sample extraction from the gel to the droplet. Comparing to Figure 3 in the manuscript, smaller electrode (1 mm in diameter) and holes (0.5 mm in diameter) are employed. Under this condition, the spatial resolution of gel ESTASI-MS is ~2 mm.

SI-4: Peptides identified from BSA digest by IEF-ESTASI-MS

MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFA
KTCVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTL
CDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLPKIETMREKVLASSAR
QRLRCASIQKFGERALKAWSVARLSQKFPKAEFVEVTKLVTDLTKVHKECCHGDLLECADDRADLAKYICDN
QDTISSKLKECCDKPLLEKSHCIAEVEKDAIPENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYEYSRRHPEY
AVSVLLRLAKEYEATLEECCAKDDPHACYSTVFDKLKHLVDEPQNLIKQNCDQFEKLGEYGFQNALIVRYTR
KVPQVSTPTLVEVSRSLGKVGTRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTESLVNRRPC
FSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQTALVELLKHKPKATEEQLKTVMENFVAFVDKCCA
ADDKEACFAVEGPKLVVSTQTALA

Scheme SI-4. Sequence of bovine serum albumin. The amino acid residues in red colour were identified by the IEF-ESTASI-MS from 4 places of the gel (cathode, anode, pH = 5.8 and pH = 6.2). The identification sequence coverage from these 4 places was found as 74%.

Table SI-4.1 Peptides identified from BSA digest by in-gel IEF-ESTASI-MS from an area of the gel close to anode. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; MSO: methionine sulfoxide; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/). The m/z is the monoisotopic peak value on the mass spectrum in figure 4 anode. These peaks can be singly, doubly or triply protonated ions.

Sequence	PTMs	M.W.	m/z	p*I*
AEFVEVTK		921.48	922.75	4.5
QNCDQFEK		1010.41	506.36, 1011.70	4.4
LVNELTEFAK		1162.62	1164.00	4.5
ETYGDMADCCEK		1363.47	455.83	3.9
YICDNQDTISSK		1385.61	463.12	4.2
TVMENFVAFVDK		1398.69	1400.00	4.4
ETYGDMADCCEK	Cys_CAM	1420.47	711.04, 1421.00	4.2
YICDNQDTISSK	Cys_CAM	1442.61	722.46, 1443.90	4.2
EYEATLEECCAK	Cys_CAM	1444.56	482.41	4.1
ETYGDMADCCEK	2Cys_CAM, MSO	1493.47	498.80	3.9
LKPDPNTLCDEFK	Cys_CAM	1575.74	526.34, 1577.00	4.6
ECCHGDLLECADDR		1577.59	789.97	4.1
ECCHGDLLECADDR	2Cys_CAM	1691.59	1693.00	4.1
YNGVFQECCQAEDK	2Cys_CAM	1746.65	583.31	4.1
CCAADDKEACFAVEGPK		1755.73	1757.00	4.3
VASLRETYGDMADCCEK		1889.80	630.78	4.3
LFTFHADICTLPDTEK	Cys_CAM	1906.89	636.73, 1908.1	4.5
LKPDPNTLCDEFKADEK		1961.94	655.04	4.4
VASLRETYGDMADCCEK	2Cys_CAM, MSO	2019.80	674.31	4.3
ECCHGDLLECADDRADLAK	3Cys_CAM	2246.87	1124.56	4.4
VASLRETYGDMADCCEKQEPER	MSO	2545.09	849.45	4.4
EYEATLEECCAKDDPHACYSTVFDK		2866.18	956.42	4.2
EYEATLEECCAKDDPHACYSTVFDKLK	2Cys_CAM	3221.36	1074.9	4.4
LAKEYEATLEECCAKDDPHACYSTVFDK	2Cys_CAM	3292.39	1647.39	4.4
ECCHGDLLECADDRADLAKYICDNQDTISSK		3443.47	1722.90	4.2
SHCIAEVEKDAIPENLPPLTADFAEDKDVCK	2Cys_CAM	3510.62	1171.33	4.3
ECCHGDLLECADDRADLAKYICDNQDTISSKLK	4Cys_CAM	3912.65	1305.30	4.4
ETYGDMADCCEKQEPERNECFLSHKDDSPDLPK	2Cys_CAM	3942.60	1972.50	4.3

Table SI-4.2 Peptides identified from BSA digest by in-gel ESTASI-MS from an area of the gel with pH around 5.8. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/). The m/z is the monoisotopic peak value on the mass spectrum in figure 4 pH = 5.8. These peaks can be singly or doubly protonated ions. The peptides in this table were identified by several ESTASI-MS from the same place of the same gel, while the figure 4 pH = 5.8 shows only one of these mass spectra.

Sequence	PTMs	M.W.	m/z	p*I*
LVTDLTK		788.47	789.24	5.8
YLYEIAR		926.49	927.7	6.0
LVVSTQTALA		1001.58	1002.70	5.5
QTALVELLK		1013.62	1015.00	6.0
CCTKPESER		1051.45	1052.80	6.1
CCTESLVNR	2Cys_CAM	1137.45	1138.90	6.0
DVCKNYQEAK	Cys_CAM	1253.55	627.80	6.1
LKECCDKPLLEK		1417.74	710.00, 1419.10	6.2
LGEYGFQNALIVR		1478.80	740.52, 1480.10	6.0
LKECCDKPLLEK	2Cys_CAM	1531.78	767.04	6.2
YICDNQDTISSKLK		1626.80		6.0
RPCFSALTPDETYVPK		1822.90	912.16, 1824.05	6.1
RPCFSALTPDETYVPK	Cys_CAM	1879.92	941.24, 1881.10	6.1

Table SI-4.3 Peptides identified from BSA digest by in-gel ESTASI-MS from an area of the gel with pH around 6.2. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/). The m/z is the monoisotopic peak value on the mass spectrum in figure 4 pH = 6.2. These peaks can be singly or doubly protonated ions. The peptides in this table were identified by several ESTASI-MS from the same place of the same gel, while the figure 4 pH = 6.2 shows only one of these mass spectra.

Sequence	PTMs	M.W.	m/z	p*I*
IETMR		648.33	649.51	6.0
TPVSEK		659.36	660.52	5.7
YLYEIAR		926.49	927.36	6.0
QTALVELLK		1013.62	1014.80	6.0
CCTESLVNR		1023.46	1024.68	6.0
CCTKPESER		1051.45	1052.73	6.1
CCTKPESER	Cys_CAM	1108.47		6.1
CCTESLVNR	2Cys_CAM	1137.50	1138.61	6.0
CCTKPESER	2Cys_CAM	1165.49	1166.68	6.1
DVCKNYQEAK	Cys_CAM	1253.58	627.86, 1254.70	6.1
LKECCDKPLLEK		1417.74	710.00, 1418.90	6.2
LGEYGFQNALIVR		1478.80	740.50, 1480.00	6.0
VPQVSTPTLVEVSR		1510.84	756.49, 1512.11	6.0
LKECCDKPLLEK	2Cys_CAM	1531.78	767.01, 1532.93	6.2
YICDNQDTISSKLK	Cys_CAM	1683.82		6.0
DAFLGSFLYEYSRR		1722.84		6.1
RPCFSALTPDETYVPK		1822.90		6.1
RPCFSALTPDETYVPK	Cys_CAM	1879.92	941.08, 1881.13	6.1
NYQEAKDAFLGSFLYEYSRR		2456.18	1229.19	6.2

Table SI-4.4 Peptides identified from BSA digest by in-gel ESTASI-MS from an area of the gel close to cathode. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/). The m/z is the monoisotopic peak value on the mass spectrum in figure 4 cathode. These peaks can be singly or doubly protonated ions. The peptides in this table were identified by several ESTASI-MS from the same place of the same gel, while the figure 4 cathode shows only one of these mass spectra.

Sequence	PTMs	M.W.	m/z	p*I*
DTHK		499.25		6.74
GACLLPK	Cys_CAM	757.42	758.61	8.22
LCVLHEK	Cys_CAM	897.48	898.65	6.74
ALKAWSVAR		1000.59	1001.73	11
GVFRRDTHK		1114.61		10.84
CASIQKFGER		1137.57		8.22
KQTALVELLK		1141.71	1142.90	8.59
QRLRCASIQK		1201.68		10.86
HPEYAVSVLLR		1282.71	642.5, 1283.90	6.75
HKPKATEEQLK		1307.73	954.95	8.51
QIKKQTALVELLK		1510.95		9.7
AWSVARLSQKFPK		1516.86		11.17
KVPQVSTPTLVEVSR		1638.94	820.6, 1640.10	8.75

SI-5: Peptides identified from BSA digest by IEF-MALDI-MS

MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFA
KTCVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTL
CDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLPKIETMREKVLASSAR
QRLRCASIQKFGERALKAWSVARLSQKFPKAEFVEVTKLVTDLTKVHKECCHGDLLECADDRADLAKYICDN
QDTISSKLKECCDKPLLEKSHCIAEVEKDAIPENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYEYSRRHPEY
AVSVLLRLAKEYEATLEECCAKDDPHACYSTVFDKLKHLVDEPQNLIKQNCDQFEKLGEYGFQNALIVRYTR
KVPQVSTPTLVEVSRSLGKVGTRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTESLVNRRPC
FSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQTALVELLKHKPKATEEQLKTVMENFVAFVDKCCA
ADDKEACFAVEGPKLVVSTQTALA

Scheme SI-5. Sequence of bovine serum albumin. The amino acid residues in red colour were identified by the IEF-MALDI-MS from 4 places of the gel (cathode, anode, pH = 5.8 and pH = 6.2). The identification sequence coverage from these 4 places was found as 47%.

Table SI-5.1 Peptides identified from BSA digest by IEF-MALDI-MS from an area of the gel close to anode. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/).

Sequence	PTMs	M.W.	p*I*
AFDEK		608.29	4.4
QEPER		657.32	4.5
ATEEQLK		817.43	4.5
DDSPDLPK		885.42	3.9
AEFVEVTK		921.49	4.5
QNCDQFEK	Cys_CAM	1067.44	4.4
LVNELTEFAK		1162.63	4.5
TVMENFVAFVDK		1398.69	4.4
YICDNQDTISSK	Cys_CAM	1442.64	4.2
EYEATLEECCAK	2Cys_CAM	1501.61	4.1
ECCHGDLLECADDR	3Cys_CAM	1748.66	4.1
LKPDPNTLCDEFKADEK	Cys_CAM	2018.97	4.4

Table SI-5.2 Peptides identified from BSA digest by IEF-MALDI-MS from an area of the gel around pH 5.8. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/).

Sequence	PTMs	M.W.	p*I*
LVTDLTK		788.47	5.8
LVVSTQTALA		1001.58	5.5
QTALVELLK		1013.62	6.0
CCTESLVNR	2Cys_CAM	1137.49	6.0
FKDLGEEHFK		1248.62	5.5
LGEYGFQNALIVR		1478.79	6.0
VPQVSTPTLVEVSR		1510.84	6.0
LKECCDKPLLEK	2Cys_CAM	1531.78	6.2
RPCFSALTPDETYVPK	Cys_CAM	1879.92	6.1

Table SI-5.3 Peptides identified from BSA digest by IEF-MALDI-MS from an area of the gel around pH 6.2. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/).

Sequence	PTMs	M.W.	p*I*
NYQEAK		751.36	6.0
QTALVELLK		1013.62	6.0
CCTESLVNR	2Cys_CAM	1137.49	6.0
CCTKPESER	2Cys_CAM	1165.49	6.1
LGEYGFQNALIVR		1478.79	6.0
VPQVSTPTLVEVSR		1510.84	6.0
LKECCDKPLLEK	2Cys_CAM	1531.78	6.2
RPCFSALTPDETYVPK	Cys_CAM	1879.92	6.1

Table SI-5.4 Peptides identified from BSA digest by IEF-MALDI-MS from an area of the gel around close to cathode. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/).

Sequence	PTMs	M.W.	p*I*
LCVLHEK	Cys_CAM	897.48	6.7
NECFLSHK	Cys_CAM	1033.48	6.7
DTHKSEIAHR		1192.59	6.9
HPEYAVSVLLR		1282.70	6.8
RHPEYAVSVLLR		1438.80	8.8
ALKAWSVARLSQK		1456.85	11.2
QIKKQTALVELLK		1510.94	9.7
KVPQVSTPTLVEVSR		1638.93	8.8

SI-6: Peptides identified from BSA digest by OFFGEL IEF-ESI-MS

MKWVTFISLLLLFSSAYSRGVFRRDTHKSEIAHRFKDLGEEHFKGLVLIAFSQYLQQCPFDEHVKLVNELTEFA
KTCVADESHAGCEKSLHTLFGDELCKVASLRETYGDMADCCEKQEPERNECFLSHKDDSPDLPKLKPDPNTL
CDEFKADEKKFWGKYLYEIARRHPYFYAPELLYYANKYNGVFQECCQAEDKGACLLPKIETMREKVLASSAR
QRLRCASIQKFGERALKAWSVARLSQKFPKAEFVEVTKLVTDLTKVHKECCHGDLLECADDRADLAKYICDN
QDTISSKLKECCDKPLLEKSHCIAEVEKDAIPENLPPLTADFAEDKDVCKNYQEAKDAFLGSFLYEYSRRHPEY
AVSVLLRLAKEYEATLEECCAKDDPHACYSTVFDKLKHLVDEPQNLIKQNCDQFEKLGEYGFQNALIVRYTR
KVPQVSTPTLVEVSRSLGKVGTRCCTKPESERMPCTEDYLSLILNRLCVLHEKTPVSEKVTKCCTESLVNRRPC
FSALTPDETYVPKAFDEKLFTFHADICTLPDTEKQIKKQTALVELLKHKPKATEEQLKTVMENFVAFVDKCCA
ADDKEACFAVEGPKLVVSTQTALA

Scheme SI-6. Sequence of bovine serum albumin. The amino acid residues in red colour were identified by the OFFGEL IEF-ESI-MS from 4 places of the gel (cathode, anode, pH = 5.8 and pH = 6.2). The identification sequence coverage from these 4 places was found as 48%.

Table SI-6.1 Peptides identified from BSA digest by OFFGEL IEF-ESI-MS. The samples were taken from the well close to anode. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/).

Sequence	PTMs	M.W.	p*I*
DDSPDLPK		885.42	3.9
QNCDQFEK		1010.42	4.4
EACFAVEGPK		1049.49	4.5
QNCDQFEK	Cys_CAM	1067.44	4.4
YICDNQDTISSK	Cys_CAM	1442.64	4.2
ETYGDMADCCEK	2Cys_CAM	1477.52	3.9
EYEATLEECCAK	2Cys_CAM	1501.61	4.1
YNGVFQECCQAEDK	2Cys_CAM	1746.71	4.1
DAIPENLPPLTADFAEDK		1954.96	3.8
TVMENFVAFVDKCCAADDK	Cys_CAM	2161.96	4.2
ECCHGDLLECADDRADLAK	3Cys_CAM	2246.94	4.2
DAIPENLPPLTADFAEDKDVCK	Cys_CAM	2457.18	4.0

Table SI-6.2 Peptides identified from BSA digest by OFFGEL IEF-ESI-MS. The samples were taken from the well with pH around 5.8. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/).

Sequence	PTMs	M.W.	p*I*
IETMR		648.33	6.0
TPVSEK		659.36	5.7
NYQEAK		751.36	6.0
LVTDLTK		788.47	5.8
LVVSTQTALA		1001.58	5.5
SHCIAEVEK		1014.49	5.4
CCTESLVNR		1023.46	6.0
CCTESLVNR	2Cys_CAM	1137.50	6.0
CCTKPESER	2Cys_CAM	1165.49	6.1
FKDLGEEHFK		1248.62	5.5
LGEYGFQNALIVR		1478.80	6.0
VPQVSTPTLVEVSR		1510.84	6.0
RPCFSALTPDETYVPK	Cys_CAM	1879.92	6.1

Table SI-6.3 Peptides identified from BSA digest by OFFGEL IEF-ESI-MS. The samples were taken from the well with pH around 6.2. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/).

Sequence	PTMs	M.W.	p*I*
ADLAK		516.30	5.9
IETMR		648.33	6.0
TPVSEK		659.36	5.8
LVTDLTK		788.47	5.8
CCTESLVNR		1023.46	6.0
DVCKNYQEAK	Cys_CAM	1253.58	6.1
LGEYGFQNALIVR		1478.80	6.0
VPQVSTPTLVEVSR		1510.84	6.0
LKECCDKPLLEK	2Cys_CAM	1531.78	6.2
RPCFSALTPDETYVPK	Cys_CAM	1879.92	6.1

Table SI-6.4 Peptides identified from BSA digest by OFFGEL IEF-ESI-MS. The samples were taken from the well close to cathode. PTMs: posttranslational modifications; Cys_CAM: carbamidomethylation cysteine; M.W.: monoisotopic molecular weight. The p*I* (isoelectric point) was calculated with the Compute p*I*/Mw tool on ExPASy (http://web.expasy.org/compute_pi/).

Sequence	PTMs	M.W.	p*I*
AWSVAR		688.37	9.8
SEIAHRFK		986.54	8.5
GVFRRDTHK		1114.61	10.8
DTHKSEIAHR		1192.60	6.9
VLASSARQRLR	Cys_CAM	1255.75	12.3

SI-7: IEF-ESTASI-MS analysis of *Escherichia coli* protein extract spiked with myoglobin and cytochrome c

Escherichia coli protein extracts were mixed together with myoglobin and cytochrome c solution for IEF-ESTASI-MS analysis. The protein mixture was first separated by IEF with a 24 cm polyacrylamide gel (pH 4-7) and then analysed by ESTASI-MS. Shown as figure SI-7, the myoglobin and cytochrome c were detected from the complex protein mixture by an ion trap mass spectrometer after IEF separation. 3 mass spectra obtained from other places of the gel were also displayed. The identification of proteins in these 3 mass spectra is impossible with the low-resolution ion trap mass spectrometer and without powerful gas phase fragmentation strategy for protein ions and special data analysis software. Besides, the spatial resolution of IEF is not enough to completely fractionate all the proteins in the complex *Escherichia coli* extract. Several proteins may be present in one fraction, and each mass spectrum can be obtained from several proteins. However, it is clear that the proteins separated from a complex sample can be *in-situ* ionised from the gel. With the development of other technologies, such as high resolution MS, gas phase fragmentation of protein ions and data analysis software, the gel-ESTASI-MS has an attractive application potential in top-down proteomics.

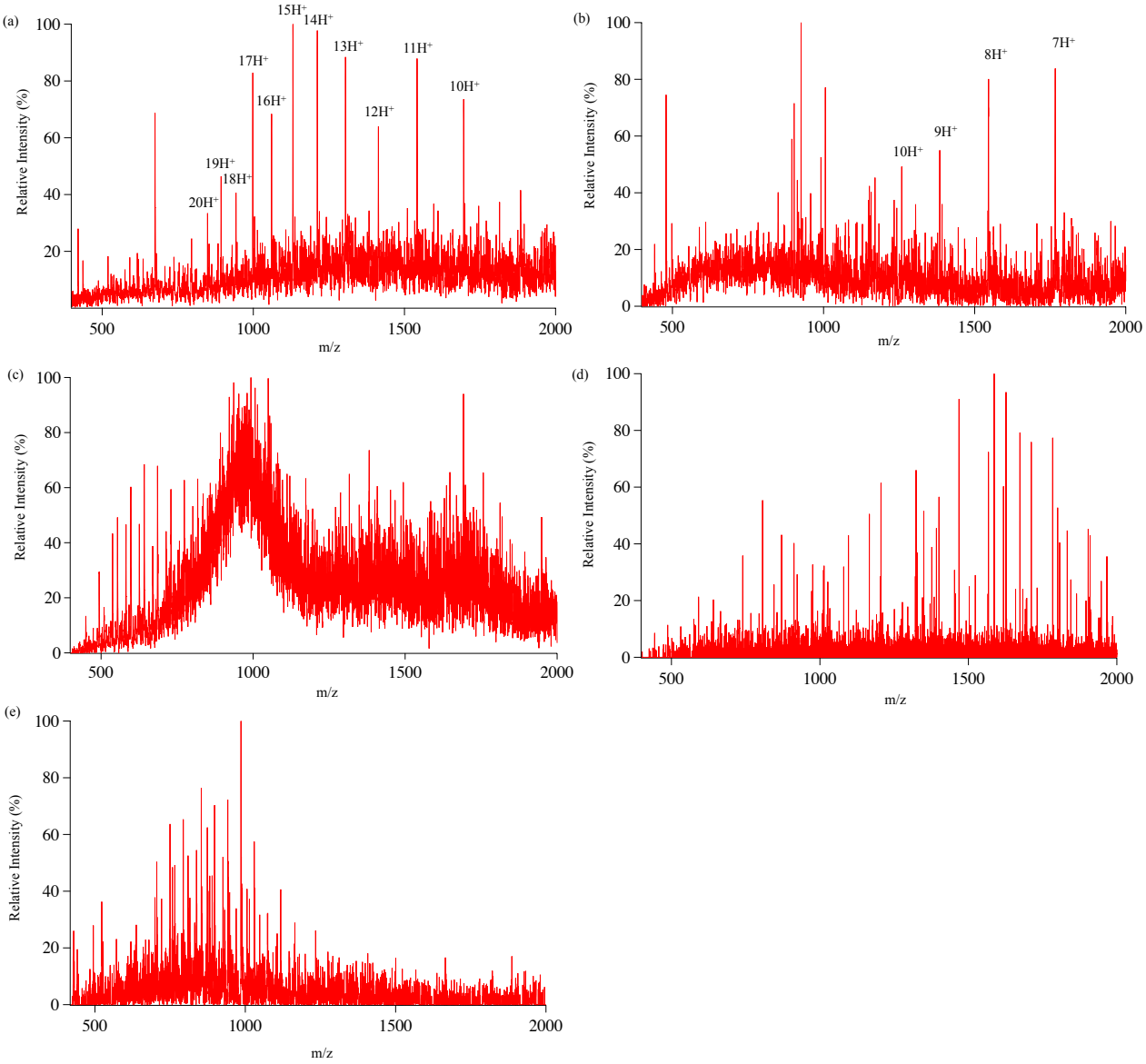


Figure SI-7. ESTASI-MS analysis of labelled *E.coli* extract samples separated by IEF: (a) fraction containing myoglobin, (b) fraction containing cytochrome c, (c) fraction at pH = 6.6, (d) fraction at pH = 5.1 and (e) fraction at pH = 4.4. The label shows the charge states.